Reproductive BioMedicine Online (2014) **29**, 729–736

www.sciencedirect.com
www.rbmonline.com

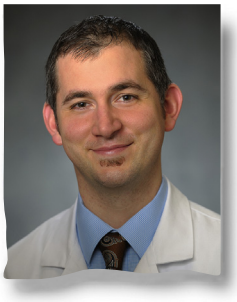
ARTICLE

Computer-automated time-lapse analysis results correlate with embryo implantation and clinical pregnancy: A blinded, multi-centre study

Matthew D VerMilyea [a], Lei Tan [b], Joshua T Anthony [a], Joe Conaghan [c], Kristen Ivani [d], Marina Gvakharia [e], Robert Boostanfar [f], Valerie L Baker [g], Vaishali Suraj [b], Alice A Chen [b], Monica Mainigi [a], Christos Coutifaris [a], Shehua Shen [b,*]

[a] *Penn Fertility Care, University of Pennsylvania, 3701 Market St. Suite #800, Philadelphia, PA 19104, USA*; [b] *Auxogyn, Inc, 1490 O'Brien Drive, Suite A, Menlo Park, CA 94025, USA*; [c] *Pacific Fertility Center, 55 Francisco Street, Fifth Floor, San Francisco, CA 94133, USA*; [d] *Reproductive Science Center of the Bay Area, 3160 Crow Canyon Road, San Ramon, CA 94583, USA*; [e] *Palo Alto Medical Foundation Fertility, Fertility Physicians of Northern California, 2581 Samaritan Drive, San Jose, CA 95124, USA*; [f] *HRC Fertility, 15503 Ventura Blvd, Suite 200, Encino, CA 91436, USA*; [g] *Stanford Fertility & Reproductive Medicine Center, 900 Welch Road, Suite 350, Palo Alto, CA 94304, USA*

* Corresponding author. *E-mail address:* sshen@auxogyn.com (S Shen).



Matthew VerMilyea, PhD, HCLD, is Director of Assisted Reproductive Technologies and Andrology Laboratories at the University of Pennsylvania, USA, and Scientific Director of Fertility Associates, New Zealand. He graduated from Cornell University, and received training in human embryology at Shady Grove Fertility, USA. He then obtained his PhD in Epigenetics from the University of Birmingham, UK. Matthew received a postdoctoral fellowship from the Japanese Society for the Promotion of Science for mammalian molecular embryology research at the RIKEN Institute, Kobe. Research interests include non-invasive and uninterrupted embryo culture protocols, time-lapse video imagery, micro-environment culture settings and early embryo histone modifications.

Abstract Computer-automated time-lapse analysis has been shown to improve embryo selection by providing quantitative and objective information to supplement traditional morphology. In this multi-centre study, the relationship between such computer-derived outputs (High, Medium, Low scores), embryo implantation and clinical pregnancy were examined. Data were collected from six clinics, including 205 patients whose embryos were imaged by the Eeva™ System. The Eeva scores were blinded and not considered during embryo selection. Embryos with High and Medium scores had significantly higher implantation rates than those with Low scores (37% and 35% versus 15%; $P < 0.0001$; $P = 0.0004$). Similar trends in implantation rates were observed in different IVF centres each using their own protocols. Further analysis revealed that patients with at least one High embryo transferred had significantly higher clinical pregnancy rates than those with only Low embryos transferred (51% versus 34%; $P = 0.02$), although patients' clinical characteristics across groups were comparable. These data, together with previous research and clinical studies, confirm

http://dx.doi.org/10.1016/j.rbmo.2014.09.005

that computer-automated Eeva scores provide valuable information, which may improve the clinical outcome of IVF procedures and ultimately facilitate the trend of single embryo selection.

KEYWORDS: Eeva Test, embryo selection, implantation, in-vitro fertilization, pregnancy, time-lapse

Introduction

Elective single-embryo transfer (SET) is increasingly promoted in IVF clinical practice as an approach to help reduce multiple births (Johnston et al., 2014), which have been associated with higher risk of adverse perinatal and maternal outcomes (Gerris et al., 1999, 2002; van Montfoort et al., 2005; Vilska et al., 1999), in addition to increased financial cost (Bromer et al., 2011; De Sutter et al., 2002). The challenge of using elelctive SET in clinical practice, however, lies in ensuring that the selection of a single embryo for transfer results in the same pregnancy rates compared with double-embryo transfer or multiple-embryo transfer. Currently, morphology based embryo selection is the most widely used method in clinical practice (Abeyta and Behr, 2014); however, snapshot morphological assessment alone has had limited success owing to dynamic changes in embryo morphology over time (Montag et al., 2011). To achieve high pregnancy rates while reducing multiple embryo transfer, improved embryo selection methods are needed.

Time-lapse imaging is an emerging, non-invasive technology that allows for the identification of promising kinetic parameters that may improve the success rates of embryo selection (Aparicio et al., 2013; Chen et al., 2013; Kirkegaard et al., 2012; Wong et al., 2013). Recently, a number of embryo selection methods based on time-lapse imaging have been reported. Meseguer et al. (2011) used known implantation data from day 3 embryo transfers, and developed a hierarchical multivariable embryo selection method that categorizes embryos into 10 grades using morphology and three timing parameters: t5 (time to 5-cell), s2 (synchrony in divisions from two-cell to four-cell) and cc2 (duration of second cell cycle). Later, the same group applied this model retrospectively to examine patient pregnancy as an end-point, and found that the classification results correlated to pregnancy rates (Meseguer et al., 2012). More recently, embryo selection models have been developed using aneuploidy data obtained after blastomere or trophectoderm biopsy. Campbell et al. (2013a) developed an aneuploidy risk classification model for blastocyst stage embryos based on the start time of blastulation and time to the fully expanded blastocyst stage. The same group later evaluated the model against blastocyst implantation in a retrospective study, and found the classification results correlated to probability of implantation (Campbell et al., 2013b). Basile *et al.* (2014) published an embryo selection method, also based on embryo developmental kinetics, which ranks the probability of day 3 embryos being chromosomally normal. These studies are valuable towards evaluating embryo selection models in appropriately designed randomized controlled trials. To the best of our knowledge, however, the methods above have not yet been validated using prospective independent clinical data and the equivalent end-points that the models were developed upon.

Recently, Conaghan et al. (2013) published an embryo classification model that, when used adjunctively with traditional morphology, can predict embryo development potential at the cleavage stage (the Eeva™ Test). The classification model was based on two cell division timing parameters, P2 (time between the first and second mitosis, or duration of the 2-cell stage) and P3 (time between the second and third mitosis, or duration of the three-cell stage), both of which have been shown in different studies to consistently correlate to embryo development, implantation potential, or both (Chavez et al., 2012; Cruz et al., 2012; Hlinka et al., 2012; Meseguer et al., 2011; Rubio et al., 2012; Wong et al., 2010). Importantly, P2 and P3 could be automatically extracted from time-lapse video recordings by state-of-the-art computer vision software that helps reduce the need for manual video annotation (Conaghan et al., 2013). Conaghan et al. (2013) validated the performance of the model in several unique ways. First, an independent dataset was used to validate the model performance. Second, the dataset was probed for the same end-point that the model was originally established upon. Finally, the model was tested adjunctively with traditional morphology and shown to significantly improve a group of embryologists' abilities to identify embryos with higher developmental potential. To further assess the clinical value of the model, the aim of this study was to examine if computer-automated prediction scores for embryo development potential also correlate to embryo implantation, patient pregnancy, or both.

Materials and methods

Data included in this retrospective study were collected from six IVF clinics in the USA. Institutional Review Board approval was obtained at each site (IRB reference number and date of approval for each site: University of Pennsylvania: number 817180, 16 January 2013; Pacific Fertility Center: number 1122828, 29 December 2011; HRC Fertility: number 1128064, 27 September 2011; Reproductive Science Centre: number 1122464, 22 December 2011; Fertility Physicians of Northern California: number 1122761, 29 December 2011; Stanford University: number 20335, 11 January 2011). Patients undergoing assisted reproduction treatment between June 2011 and December 2013 were approached and consented to have their embryos imaged. Patients underwent ovarian stimulation according to guidelines of each clinic, where protocols included agonist luteal phase, agonist microdose flare and antagonist suppression. On the day of oocyte retrieval (day 0), oocytes were fertilized using conventional insemination protocols or by intracytoplasmic sperm injection. Immediately after fertilization assessment, normally fertilized oocytes (2PNs) were placed into a multiwell Eeva Dish. The microwell format of the petri-dish holds

individual embryos separately but in close proximity under a shared media droplet, whereas reference labels provide visual orientation of each embryo's specific location within the dish array. Throughout embryo culture, each clinical site used its own laboratory protocols, including their standard culture media, protein supplementation, media exchange protocol (if necessary) and incubation environment.

Embryos were imaged using the Eeva System (Auxogyn, Inc., Menlo Park, CA, USA), a time-lapse enabled system that generates automated prediction scores to predict embryo development in adjunct with traditional morphology. Briefly, the Eeva System includes microscopes that fit into a standard incubator and use dark field imaging to capture high resolution, single-plane images of embryos housed in a multi-well Eeva Dish, in 5-min intervals. To maintain a continuous and uninterrupted imaging process from day 1 to day 3, no media changes or excursions from the incubator were permitted after day 1 and before day 3. Embryo morphology assessment on different culture days, embryo grading, selection and transfer were carried out according to the standard operating procedures of each individual clinic. As this was a blinded non-selection study, embryos that were selected for transfer or for freezing were based on traditional morphology grading only according to each clinic's protocol. Clinical pregnancy and successful implantation were confirmed by ultrasound showing evidence of intrauterine fetal heart beat at about 6–8 weeks' gestational age. Implantation rate was defined as the number of gestational sacs with fetal heart beat divided by the number of embryos transferred. Clinical pregnancy rate was defined as the number of patients who had visible fetal heartbeat at 6–8 weeks' gestational age divided by the total number of patients.

Image data were analysed by the Eeva System's computer vision software, which automatically extracts P2 and P3 timings for each embryo, feeds the timings into an established statistical classification model, and generates a predictive score of whether the embryo has a 'High' or 'Low' probability of developing to a usable blastocyst (Conaghan et al., 2013). In this retrospective analysis, two different versions of the Eeva outputs were studied: the original two-category output, which gives an Eeva High (P2: $9.33 \leq P2 \leq 11.45$ h; P3: $0 \leq P3 \leq 1.73$ h) or Eeva Low (P2, P3, or both, out of the Eeva High window) score (Conaghan et al., 2013); and a new, further sub-divided, three-category output which gives an Eeva High ($9.33 \leq P2 \leq 11.45$ h; $0 \leq P3 \leq 1.73$ h), Eeva Medium (not Eeva High: $9.33 \leq P2 \leq 12.65$ h; P3: $0 \leq P3 \leq 4$ h) or Eeva Low (P2, P3, or both, out of the Eeva High and Eeva Medium window) score. Rates of implantation and clinical pregnancy could be calculated for all single embryo transfers and, in cases of multiple embryo transfers, when all transferred embryos were of the same Eeva score.

Statistical data analysis was carried outu using SAS 9.3 (SAS Institute Inc., North Carolina, USA). Student's *t*-test, chi-squared test, Fisher's exact or one-way analysis of variance was used to calculate *P*-values. A comparison with $P < 0.05$ was considered to have statistically significant difference.

Results

This retrospective analysis included 205 patients whose transferred embryos had P2 and P3 timings automatically extracted by the computer programme, to investigate a possible association between the Eeva scores and rates of embryo implantation and clinical pregnancy. The patient clinical characteristics are described in **Table 1**. From these 205 patients, a total of 375 embryos were transferred, resulting in an overall 29% (110/375) implantation rate and 45% (93/205) clinical pregnancy rate. Of the embryos that were transferred, 331 embryos had implantation status confirmed; implantation status of 44 embryos could not be determined owing to transfer of multiple embryos with different Eeva scores; therefore, they were excluded from data analysis.

Two-category (High/Low) output and rates of embryo implantation and clinical pregnancy

To examine the association between the Eeva two-category output (High/Low) and embryo implantation, we analysed the implantation rates for transferred embryos with known implantation ($n = 331$ embryos). The overall known implantation rate for this dataset was 27% (91/331) and was calculated using only those patients for whom implantations could be matched with specific embryo(s) transferred. Known implantation rate is typically lower than implantation rates that are calculated when all patients are included (Rubio et al., 2012), because some of the cases with partial implantation are excluded and therefore non-implanted embryos are over-represented. We found that High embryos had a higher probability of successful implantation (37%, 41/111) than Low embryos (23%, 50/220) (**Figure 1**), and the difference in implantation rates was statistically significant ($P = 0.003$).

To examine the association between the Eeva two-category output (High and Low) and clinical pregnancy rates, 205 patients were divided into two groups: patients with at least one High embryo transferred and those with no High embryo transferred. The patients' clinical characteristics for the two groups were compared, including egg age, number of eggs retrieved, number of 2PNs on day 1 and number of embryos transferred. No statistically significant difference were found for any of the clinical characteristics assessed (**Table 2**). Patients with at least one High embryo, transferred, however, had statistically significantly higher clinical pregnancy rates than those with no High embryos transferred (51% versus 39%; $P = 0.04$) (**Table 2**).

Three-category (High, Medium and Low) output and rates of embryo implantation and clinical pregnancy

As 50 out of 205 patients (24%) only had Low embryos available, we hypothesized that having a three-category Eeva output that further separated the Low category into a Medium and a Low subcategory may help embryo selection for these patients. With the three-category output, 155 out of 205 (76%) patients had High embryos, 31 out of 205 patients (15%) had no High but had Medium, Low embryos, or both, and 19 out of 205 of patients (9%) had only Low embryos available. Indeed, after introducing the Medium category, the frequency of patients who had only Low embryos dropped from 24% to 9%.

We next examined if the three-category output (High, Medium and Low) relates to embryo implantation. High

Table 1 Clinical characteristics for patients included in this study.

Total number of patients		205
Total number of eggs		3523
Total number of 2PNs		2142
Egg age (years) (mean ± SD)		33.1 ± 5.1
Recipient age (years) (mean ± SD)		35.7 ± 5.2
Cycle type	Patient using own eggs *n* (%)	176/205 (86)
	Oocyte donor *n* (%)	29/205 (14)
Reason for assisted reproductive technology	Male infertility *n* (%)	39/205 (19)
	History of endometriosis *n* (%)	5/205 (2)
	Ovulation disorders *n* (%)	14/205 (7)
	Diminished ovarian reserve *n* (%)	33/205 (16)
	Tubal factor *n* (%)	5/205 (2)
	Uterine *n* (%)	2/205 (1)
	Unexplained/other *n* (%)	53/205 (26)
	Multiple reasons *n* (%)	54/205 (26)
Stimulation protocol	Agonist luteal phase *n* (%)	55/205 (27)
	Agonist micro-dose flare *n* (%)	11/205 (5)
	Antagonist suppression *n* (%)	102/205 (50)
	Other *n* (%)	37/205 (18)
Egg retrieval count	Number of eggs (mean ± SD)	17.4 ± 8.5
Method of insemination	Intracytoplasmic sperm injection *n* (%)	119/205 (58)
	IVF *n* (%)	76/205 (37)
	Both *n* (%)	3/205 (1)
	Unknown *n* (%)	7/205 (3)
Fertilization count (mean ± SD)	Number of 2PNs (mean ± SD)	10.6 ± 5.5

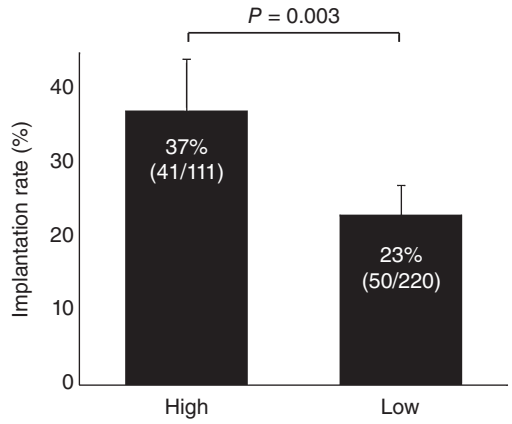


Figure 1 Implantation rates for Eeva two-category High versus Low scored embryos. Error bars represent 95% upper confidence limit. The difference in implantation rates between Eeva High versus Eeva Low embryos is statistically significant: $P = 0.003$.

embryos had the highest likelihood of implantation (37%, 41/111), followed by Medium (35%, 29/83), and Low (15%, 21/137) (Figure 2). The difference in implantation rates between high compared with low embryos and medium compared with low embryos was statistically significant ($P < 0.0001$; $P = 0.0004$, respectively), suggesting that implantation rates are also associated with an Eeva three-category output.

To analyse the relationship between the three-category output and clinical pregnancy rates, patients were divided into three groups: at least one High embryo transferred; no High but at least one Medium transferred; no High or Medium transferred. Similar to the two-category breakdown, the patients' clinical characteristics across three groups, measured by egg age, number of eggs, number of embryos and number of embryos transferred, were statistically equivalent (tested by one-way analysis of variance). Patients with one or more High embryos transferred, however, had statistically significantly higher clinical pregnancy rates than those with no High or Medium embryos transferred (51% versys 34%; $P = 0.02$) (Table 3). These results show that the difference in clinical pregnancy rates between High and Low patient groups can be augmented when a Medium category is introduced.

Eeva scores and embryo implantation in different clinics

Finally, the association between Eeva scores and embryo implantation in the three clinics that contributed most of the data (defined by having at least 50 or more transferred embryos with known implantation) were analysed. The three clinics followed their own standard procedure for embryo culture and embryo selection for transfer, where variations included IVF versus intracytoplasmic sperm injection, oxygen concentration at 5% versus 20%, VitroLife versus Sage versus Irvine Scientific culture media, and sequential versus single-step media culture protocols. Different overall known implantation rates were observed for each site (site A: 25%; site B: 32%; site C: 23%). The relationship between Eeva scores and implantation, however, was observed across all three sites: High embryos consistently had better implantation rates than

Table 2 Clinical pregnancy rates and clinical characteristics for patients whose embryos were assessed by the Eeva System and had at least one Eeva High embryo transferred or no Eeva High embryos transferred.

Patient group	*Numbr of patients*	*Egg age*	*Numbr of eggs retrieved*	*Number of 2PN*	*Number of embryos transferred*	*Pregnancy rate*
At least one Eeva High transferred	105	33.0 ± 4.9	17.7 ± 8.4	11.0 ± 5.5	1.8 ± 0.8	51% (54/105)
No Eeva High embryos transferred	100	33.2 ± 5.3	17.0 ± 8.7	10.1 ± 5.5	1.8 ± 0.7	39% (39/100)[a]

[a]$P = 0.04$.

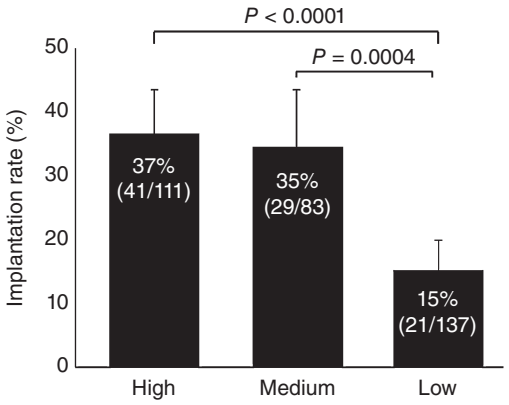


Figure 2 Implantation rates for Eeva three-category High, Medium and Low scored embryos. Error bars represent 95% upper confidence limit. Implantation rates between High versus Low and Medium versus Low were significantly different ($P < 0.0001$ and $P = 0.0004$, respectively).

Low embryos, and the difference in implantation rates between these two categories was statistically significant for all three sites (sites A and B: $P = 0.02$; site C: $P = 0.002$) (**Figure 3**).

Discussion

In this blinded multi-centre study, our results indicate that computer-automated time-lapse analyses data are related to embryo implantation and clinical pregnancy. The three-category output by computer automated analyses further set apart the difference in implantation rates between High and Low scored embryos. Furthermore, the association between the computer-generated predictive scores and embryo implantation was observed for different IVF centres using their own embryo culture protocols.

Timing parameters, P2 and P3, have been shown in different studies to consistently correlate to embryo development, implantation potential, or both (Chen et al., 2013). Previously, it was demonstrated that the Eeva two-category scores based on P2 and P3 timings could automatically predict an embryo's potential to develop into a usable blastocyst (Conaghan et al., 2013). In the present study, this automated prediction model is also shown to relate to embryo implantation and clinical pregnancy. Meseguer et al. (2011) previously developed a time-lapse based embryo classification model, using implantation as an end-point, which similarly included P2 and P3, although both timings required manual recording and input by human observers. The parameters P2 and P3 have since been evaluated in multiple studies from various clinics. Most of these studies have confirmed statistically significant correlations between manual assessment of P2, P3, or both, and embryo developmental potential, implantation, or both (Chavez et al., 2012; Cruz et al., 2012; Dal Canto et al., 2012; Hlinka et al., 2012; Kirkegaard et al., 2013b; Meseguer et al., 2011; Rubio et al., 2012; Wong et al., 2010). Two studies did not find correlations between P2, P3 and embryo quality (Campbell et al., 2013a; Chamayou et al., 2013); however, both of these studies evaluated P2 and P3 for a small subset of embryos (e.g. transferred blastocysts or biopsied blastocysts), instead of a broader, more generalizable embryo population. In addition to blastocyst development and embryo implantation, other developmental biological research investigating P2 and P3 have found these timings to be reflective of euploidy at the cleavage-stage (Chavez et al., 2012) and gene expression (Wong et al., 2010). Despite all possible variations in published studies, P2 and P3 still stand out as two parameters that are repeatedly reported by independent clinics and various studies to be indicative of embryo viability, compared with other parameters published (Kaser and Racowsky, 2014). As reproducibility is one of the main tenets of the scientific method, our results, together with previous and emerging published literature, demonstrate the predictive value of early time-lapse parameters such as P2 and P3 in differentiating embryos with higher developmental potential during embryo selection.

Two types of Eeva scores were analysed (i.e. two-category High and Low, and three-category High, Medium and Low). Both types of results were related to rates of embryo implantation and clinical pregnancy. The introduction of the Medium category may improve clinical utility by further discriminating the embryos that fall into the two-category low group. At the patient level, adding Medium reduced the percentage of patients with only Low embryos to transfer from 24% to 9%. At the embryo level, separating medium embryos from the two-category low group helped further differentiate the implantation potential within the Low embryos, which comprised 66% of all transferred embryos in this blinded study. In short, with the addition of a Medium category to the model, embryos chosen for transfer could be further differentiated by their implantation potential. 'Low' potential embryos were not expected to have 'no potential' as a result of adding a medium category to the model. Our results showed that embryos with High, Medium or Low Eeva scores resulted in relatively higher, medium and lower implantation and pregnancy rates, and support the conclusion that the

Table 3 Clinical pregnancy rates and clinical characteristics for patients whose embryos were assessed by the Eeva System and had at least one Eeva High embryo transferred; at least one Eeva Medium embryo transferred; or no Eeva High/Medium transferred.

Patient group	*Number of patients*	*Egg age*	*Number of eggs retrieved*	*Number of 2PN*	*Number of embryos transferred*	*Pregnancy rate*
At least one Eeva High transferred	105	33.0 ± 4.9	17.7 ± 8.4	11.0 ± 5.5	1.8 ± 0.8	51% (54/105)[a]
No Eeva High and at least one Eeva Medium transferred	53	33.4 ± 5.3	16.0 ± 7.4	10.5 ± 5.0	1.8 ± 0.7	43% (23/53)
No Eeva High or Medium transferred	47	32.9 ± 5.3	18.2 ± 10	9.6 ± 6.2	1.8 ± 0.7	34% (16/47)[a]

[a]For 'at least one Eeva High transferred' versus 'no Eeva High or Medium transferred'; $P = 0.02$.

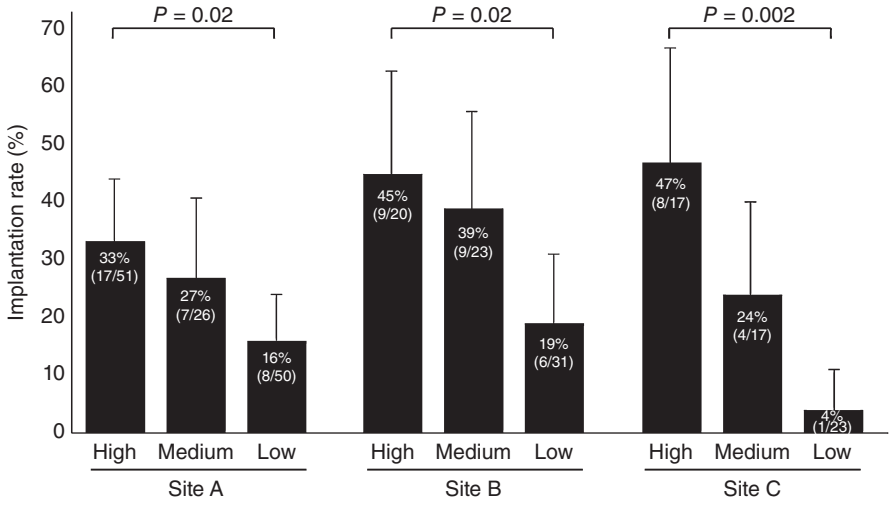


Figure 3 Implantation rates for embryos with Eeva High, Medium and Low scores from three clinical sites with at least 50 embryos of known implantation data per site. Error bars represent 95% upper confidence limit. For all three sites, the difference in implantation rates between Eeva High versus Low embryos is statistically significant (chi-squared test): $P = 0.02$ (site A); $P = 0.02$ (site B); $P = 0.002$ (site C).

Eeva information may be added to morphology to increase the chance of selecting the embryo with higher viability. As the use of time-lapse in the clinical setting increases and more imaging data become available, advanced data mining methods could be used to expand upon the current model, and future models could be developed with more categories to allow further refined differentiation among embryos non-invasively. Ultimately, a model that ranks embryos robustly for individual patients may aid in further adoption of elective SET while minimizing patients' time-to-pregnancy.

In this study, composed of data from multiple clinics, computer-derived prediction results were strongly related to embryo implantation in different IVF centres, each using their own laboratory protocols, culture media, incubation environment and insemination technique. Overall, and at individual clinics, Eeva High, Medium and Low scores resulted in relatively higher, medium and lower implantation and clinical pregnancy rates. Our results suggest that a generalizable model for optimal timing of embryo development may be feasible. Three key elements may contribute to the general applicability of the model we tested. First, the prediction model was developed using data collected from a multi-centre study (Conaghan et al., 2013) so that variability in culturing conditions was incorporated during model development. Second, the model has been validated for blastocyst prediction using an independent dataset and the same end-points as those the model was developed upon (Conaghan et al., 2013). Third, a wide variety of researchers have reported that P2 and P3 timings are negligibly influenced by a variety of external factors, including culture media (Basile et al., 2013), oxygen concentration (Kirkegaard et al., 2013a), stimulation protocols (Munoz et al., 2012) and insemination method (Dal Canto et al., 2012). Interestingly, although the present study was under review, a retrospective multi-centre analysis published in a commentary examined whether the blastocyst prediction model published by Conaghan *et al.* correlates to implantation for seven independent clinics (Kirkegaard et al., 2014). The investigators reported that embryos that were manually analysed and classified as High have a statistically significantly higher implantation rate (30% relative increase) than embryos that were manually analysed and classified as Low (Kirkegaard et al., 2014). The finding from

Kirkegaard *et al.* (2014) is consistent with our conclusion that a predictive model based on P2 and P3 relates to higher compared with lower implantation, and may therefore aid in embryo assessment. As far as is known, the present study was the first evaluation of implantation for a computer-automated test that does not require manual analysis of these critical parameters. Furthermore, results are introduced for a novel Medium group, a secondary analysis of clinical pregnancy at the patient level, and evaluation of the performance of the model at individual clinics. Our work and the work of others, therefore, support the idea of a generalizable embryo assessment model. Specifically, our study uses time-lapse and clinical data from multiple IVF centres undergoing their own standard procedures for stimulation, egg retrieval, embryo culture and insemination, suggests broad applicability for a computer-automated, P2-, P3-based predictive model in diverse clinical IVF laboratories.

As this study was blinded and time-lapse scores were not considered during embryo selection, all the cases were analysed retrospectively to estimate the percentage of patients who could potentially have had a different embryo selected if time-lapse scores were available. For the two-category Eeva output, 100 patients had no High embryos selected for transfer. Thirty-three of these patients (33%, 33/100) had at least one High embryo that was not transferred, but exhibited equivalent morphology to the embryo(s) that were selected for transfer. Notably, our patient population included those who had many candidate embryos for transfer, and were therefore in particular need of embryo selection assistance. For these patients, the availability of additional high embryo(s) suggests that, with the assistance of an automated time-lapse model, alternative embryos with higher implantation potential could have been selected for transfer. Therefore, using time-lapse results during embryo selection could potentially benefit a significant number of patients; however, the specific degree of improvement in reproductive outcome must be assessed in a well-designed randomized controlled trial, the gold standard to evaluate the effectiveness and clinical implementation of any new technology. The present retrospective patient case analysis provides key information valuable for the design of a meaningful randomised controlled trial.

This study has demonstrated that prediction models based on computer-extracted P2 and P3 timings show a strong relationship with embryo implantation and clinical pregnancy. Recent research studies have focused on the identification of additional timing parameters that correlate to embryo development, chromosomal normality, embryo implantation, or both (including those that span the first cell cycle, the cleavage stage, morula to blastocyst stage, or both) (Aguilar et al., 2014; Athayde Wirka et al., 2014; Basile et al., 2014; Campbell et al., 2013a). These new parameters, if robustly demonstrated to be predictive of embryo viability, may be combined with currently validated timings to further improve embryo selection. Increasingly, time-lapse research is appropriately focused on testing the value of embryo developmental timing parameters and selection models in prospective randomized studies (Aparicio et al., 2013). These studies promise to reset the standard of embryo selection in IVF clinical practice and move the field of assisted reproductive technology towards 'one healthy baby at a time'.

Acknowledgements

The authors gratefully acknowledge the embryologists, physicians and patients who participated in this study. We also thank Roopa Murugesan, Kelly Athayde Wirka, the clinical operations team and engineering team at Auxogyn for technical support and helpful discussions.

References

Abeyta, M., Behr, B., 2014. Morphological assessment of embryo viability. Semin. Reprod. Med. 32, 114–126.

Aguilar, J., Motato, Y., Escriba, M.J., Ojeda, M., Munoz, E., Meseguer, M., 2014. The human first cell cycle: impact on implantation. Reprod. Biomed. Online 28, 475–484.

Aparicio, B., Cruz, M., Meseguer, M., 2013. Is morphokinetic analysis the answer? Reprod. Biomed. Online 27, 654–663.

Athayde Wirka, K., Chen, A.A., Conaghan, J., Ivani, K., Gvakharia, M., Behr, B., Suraj, V., Tan, L., Shen, S., 2014. Atypical embryo phenotypes identified by time-lapse microscopy: high prevalence and association with embryo development. Fertil. Steril. 101, 1637–1648, e5.

Basile, N., Morbeck, D., Garcia-Velasco, J., Bronet, F., Meseguer, M., 2013. Type of culture media does not affect embryo kinetics: a time-lapse analysis of sibling oocytes. Hum. Reprod. 28, 634–641.

Basile, N., Nogales Mdel, C., Bronet, F., Florensa, M., Riqueiros, M., Rodrigo, L., Garcia-Velasco, J., Meseguer, M., 2014. Increasing the probability of selecting chromosomally normal embryos by time-lapse morphokinetics analysis. Fertil. Steril. 101, 699–704.

Bromer, J.G., Ata, B., Seli, M., Lockwood, C.J., Seli, E., 2011. Preterm deliveries that result from multiple pregnancies associated with assisted reproductive technologies in the USA: a cost analysis. Curr. Opin. Obstet. Gynecol. 23, 168–173.

Campbell, A., Fishel, S., Bowman, N., Duffy, S., Sedler, M., Hickman, C.F., 2013a. Modelling a risk classification of aneuploidy in human embryos using non-invasive morphokinetics. Reprod. Biomed. Online 26, 477–485.

Campbell, A., Fishel, S., Bowman, N., Duffy, S., Sedler, M., Thornton, S., 2013b. Retrospective analysis of outcomes after IVF using an aneuploidy risk model derived from time-lapse imaging without PGS. Reprod. Biomed. Online 27, 140–146.

Chamayou, S., Patrizio, P., Storaci, G., Tomaselli, V., Alecci, C., Ragolia, C., Crescenzo, C., Guglielmino, A., 2013. The use of morphokinetic parameters to select all embryos with full capacity to implant. J. Assist. Reprod. Genet. 30, 703–710.

Chavez, S.L., Loewke, K.E., Han, J., Moussavi, F., Colls, P., Munne, S., Behr, B., Reijo Pera, R.A., 2012. Dynamic blastomere behaviour reflects human embryo ploidy by the four-cell stage. Nat. Commun. 3, 1251.

Chen, A.A., Tan, L., Suraj, V., Reijo Pera, R., Shen, S., 2013. Biomarkers identified with time-lapse imaging: discovery, validation, and practical application. Fertil. Steril. 99, 1035–1043.

Conaghan, J., Chen, A.A., Willman, S.P., Ivani, K., Chenette, P.E., Boostanfar, R., Baker, V.L., Adamson, G.D., Abusief, M.E., Gvakharia, M., Loewke, K.E., Shen, S., 2013. Improving embryo selection using a computer-automated time-lapse image analysis test plus day 3 morphology: results from a prospective multicenter trial. Fertil. Steril. 100, 412–419, e5.

Cruz, M., Garrido, N., Herrero, J., Perez-Cano, I., Munoz, M., Meseguer, M., 2012. Timing of cell division in human cleavage-stage embryos is linked with blastocyst formation and quality. Reprod. Biomed. Online 25, 371–381.

Dal Canto, M., Coticchio, G., Mignini Renzini, M., de Ponti, E., Novara, P.V., Brambillasca, F., Comi, R., Fadini, R., 2012. Cleavage

kinetics analysis of human embryos predicts development to blastocyst and implantation. Reprod. Biomed. Online 25, 474–480.
De Sutter, P., Gerris, J., Dhont, M., 2002. A health-economic decision-analytic model comparing double with single embryo transfer in IVF/ICSI. Hum. Reprod. 17, 2891–2896.
Gerris, J., de Neubourg, D., Mangelschots, K., van Royen, E., van de Meerssche, M., Valkenburg, M., 1999. Prevention of twin pregnancy after in-vitro fertilization or intracytoplasmic sperm injection based on strict embryo criteria: a prospective randomized clinical trial. Hum. Reprod. 14, 2581–2587.
Gerris, J., de Neubourg, D., Mangelschots, K., van Royen, E., Vercruyssen, M., Barudy-Vasquez, J., Valkenburg, M., Ryckaert, G., 2002. Elective single day 3 embryo transfer halves the twinning rate without decrease in the ongoing pregnancy rate of an IVF/ICSI programme. Hum. Reprod. 17, 2626–2631.
Hlinka, D., Kalatova, B., Uhrinova, I., Dolinska, S., Rutarova, J., Rezacova, J., Lazarovska, S., Dudas, M., 2012. Time-lapse cleavage rating predicts human embryo viability. Physiol. Res. 61, 513–525.
Johnston, J., Gusmano, M.K., Patrizio, P., 2014. Preterm births, multiples, and fertility treatment: recommendations for changes to policy and clinical practices. Fertil. Steril. 102, 36–39.
Kaser, D.J., Racowsky, C., 2014. Clinical outcomes following selection of human preimplantation embryos with time-lapse monitoring: a systematic review. Hum. Reprod. Update 20, 802–803.
Kirkegaard, K., Agerholm, I.E., Ingerslev, H.J., 2012. Time-lapse monitoring as a tool for clinical embryo assessment. Hum. Reprod. 27, 1277–1285.
Kirkegaard, K., Hindkjaer, J.J., Ingerslev, H.J., 2013a. Effect of oxygen concentration on human embryo development evaluated by time-lapse monitoring. Fertil. Steril. 99, 738–744, e4.
Kirkegaard, K., Kesmodel, U.S., Hindkjær, J.J., Ingerslev, H.J., 2013b. Time-lapse parameters as predictors of blastocyst development and pregnancy outcome in embryos from good prognosis patients: a prospective cohort study. Hum. Reprod. 28, 2643–2651.
Kirkegaard, K., Campbell, A., Agerholm, I., Bentin-Ley, U., Gabrielsen, A., Kirk, J., Sayed, S., Ingerslev, H.J., 2014. Limitations of a time-lapse blastocyst prediction model: a large multicentre outcome analysis. Reprod. Biomed. Online 29, 156–158.
Meseguer, M., Herrero, J., Tejera, A., Hilligsoe, K.M., Ramsing, N.B., Remohi, J., 2011. The use of morphokinetics as a predictor of embryo implantation. Hum. Reprod. 26, 2658–2671.
Meseguer, M., Rubio, I., Cruz, M., Basile, N., Marcos, J., Requena, A., 2012. Embryo incubation and selection in a time-lapse monitoring system improves pregnancy outcome compared with a standard incubator: a retrospective cohort study. Fertil. Steril. 98, 1481–1489, e10.
Montag, M., Liebenthron, J., Koster, M., 2011. Which morphological scoring system is relevant in human embryo development? Placenta 32, S252–S256.
Munoz, M., Cruz, M., Humaidan, P., Garrido, N., Perez-Cano, I., Meseguer, M., 2012. Dose of recombinant FSH and oestradiol concentration on day of HCG affect embryo development kinetics. Reprod. Biomed. Online 25, 382–389.
Rubio, I., Kuhlmann, R., Agerholm, I., Kirk, J., Herrero, J., Escriba, M.J., Bellver, J., Meseguer, M., 2012. Limited implantation success of direct-cleaved human zygotes: a time-lapse study. Fertil. Steril. 98, 1458–1463.
van Montfoort, A.P., Dumoulin, J.C., Land, J.A., Coonen, E., Derhaag, J.G., Evers, J.L., 2005. Elective single embryo transfer (eSET) policy in the first three IVF/ICSI treatment cycles. Hum. Reprod. 20, 433–436.
Vilska, S., Tiitinen, A., Hyden-Granskog, C., Hovatta, O., 1999. Elective transfer of one embryo results in an acceptable pregnancy rate and eliminates the risk of multiple birth. Hum. Reprod. 14, 2392–2395.
Wong, C., Loewke, K., Bossert, N., Behr, B., de Jonge, C., Baer, T., Reijo Pera, R., 2010. Non-invasive imaging of human embryos before embryonic genome activation predicts development to the blastocyst stage. Nat. Biotechnol. 28, 1115–1121.
Wong, C., Chen, A.A., Behr, B., Shen, S., 2013. Time-lapse microscopy and image analysis in basic and clinical embryo development research. Reprod. Biomed. Online 26, 120–129.

Declaration: MDV and JC are consultants of Auxogyn, a company that develops and markets the Eeva™ System. LT, VS, AAC and SS are employees and equity holders of Auxogyn.

Received 24 April 2014; refereed 8 September 2014; accepted 9 September 2014.